

OCR 기반의 물류관리 어플리케이션

OCR-based inventory management application

서민석¹·김대한¹·강혜윤¹·최동걸[†]

Minseok Seo¹, Daehan Kim¹, Hyeyoon Kang¹, Dong-geol Choi[†]

Abstract: Accurate and fast inventory management is an essential element in the modern distribution industry. However, configuring the inventory management system is very expensive, and managing the inventory with human eyes is labor-intensive and inaccurate. We propose an OCR-based inventory management system to solve these practical problems. Our OCR-based inventory management system is designed as an application so that it can be easily used in large and small-scale distribution businesses at low cost. In addition, we have developed an OCR algorithm that is robust against various noises that exist in reality so that it can be used in a practical environment. Finally, in order to find the optimal combination of text detection and text recognition in our system, various models were integrated and analyzed. The OCR-based inventory management application proposed by us solves the problem that its application was limited, such as license plate recognition, document recognition, due to the limitation that OCR is particularly vulnerable to various noises existing in reality. Based on these results, we expect our system to be used as a inventory management system in many distribution industries.

Keywords: OCR, Attention, Inventory Management, STN, Text Detection, Text Recognition

1. 서 론

현대의 유통업에서 물류 관리 시스템을 구축하는 것은 매우 중요한 문제 중 하나이다. 유통업에서 물류 관리 시스템의 정확도는 판매에 직접적인 영향이 있기 때문이다. 하지만 높은 정확도의 물류 관리 시스템을 구축하는 것은 매우 높은 비용이 소모된다. 그렇기 때문에 규모가 작은 유통매장에서는 레이저 기반의 바코드 리더기를 사용하여 물류 관리 시스템을 구축하는 것 보다, 사람이 눈으로 제품번호를 읽고 확인하는 방법을 선호한다. 하지만 이러한 방법은 노동 집약적이며 근무하는 직원에 따라서 정확도의 차이가 가변적이기 때문에 일관적인 물류 관리 시스템을 구축하기 매우 어렵다.

우리는 이러한 문제점을 해결하기 위하여 OCR(optical character recognition)기반의 물류 관리 어플리케이션을 제안한다.

일상적인 장면에서 글자를 읽는 것은 산업에서 광범위하게 응용되는 중요한 기술이다. 최근 CNN(convolutional neural network) 기반의 영상인식 기술의 발전으로 CNN기반의 OCR^[1] 시스템은 노이즈가 없는 깨끗한 문서에서는 성공적으로 적용되었지만, 대부분의 CNN 기반의 OCR 방법들은 실제 세계에서 발생하는 노이즈, 글자모양 왜곡 등 노이즈가 있는 환경에서 약하기 때문에 응용되지 못했다. 이러한 문제를 해결하기 위하여 다양한 방법들이 제안되었지만, 그 응용은 여전히 자동차 번호판인식 문서 번역 및 복사 수준에서 멈춰있다.

우리는 OCR 기술이 노이즈에 취약하다는 고질적인 문제^[2]를 해결하기 위하여 글자 부분에 랜덤한 가우시안 노이즈를 섞고 그것을 회복하는 적대적 학습^[3]을 제안하여 문제를 해결한다. 또한 노이즈에 강인한 글자 인식 모델을 물류 관리 어플리케이션으로 확장하였다.

마지막으로 우리는 기존에 존재하는 글자 찾기, 글자 인식, 후처리 방법들을 조사하고 비교한다. 이후 조사한 방법들의

Received : Feb. 22. 2016; Revised : Apr. 19. 2016; Accepted : Jul. 21. 2016

※ This project was funded by Korea Robotics Society (KROS), and is currently supported by the publication grant

1. **Master's student**, Hanbat national University Daejeon, Korea (minsseock.seo@gmail.com)

1. **Undergraduate students**, Hanbat national University Daejeon, Korea (koreadhkim96@gmail.com)

1. **Undergraduate students**, Hanbat national University Daejeon, Korea (koycc1220@gmail.com)

† **Associate Professor**, Hanbat national University Daejeon, Korea (dgchoi@hanbat.ac.kr)

조합을 실험 및 분석하여 우리의 어플리케이션에서 최적의 구조를 설계하였다. 따라서 우리의 기여는 아래와 같다.

1. 노이즈에 강인한 학습법 제안으로, 현실에서 사용할 수 있는 OCR 알고리즘 개발.
2. 규모가 작은 유통매장에서 사용할 수 있는 신뢰성 있는 물류 관리 시스템 제안.
3. 기존의 노동 집약적인 물류 관리 시스템을 개선.

2. 관련 연구

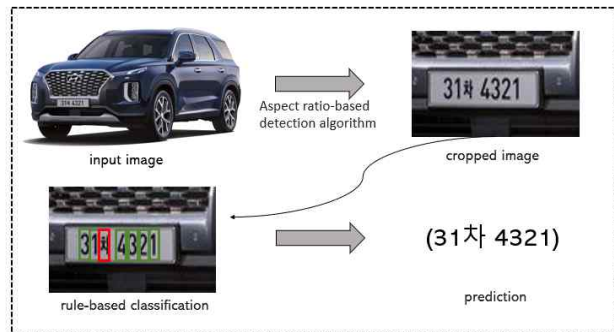
전통적인 OCR은 빛을 이용해 문자를 판독하는 장치로 종이 인쇄되거나 손으로 쓴 문자, 기호, 마크 등에 빛을 비추어 그 반사 광선을 전기 신호로 바꾸어 컴퓨터에 입력하는 장치이다. 이는 고지서, 영수증 인식에 주로 응용되어 왔다. 하지만 현재의 OCR은 디지털 문자 인식도 포함하고 있으며, CNN 기반의 OCR 기술의 발전에도 불구하고 현실에서 존재하는 노이즈에 취약하기 때문에 노이즈가 없는 영수증, 문서, 번호판 인식에 주로 사용되고 있다.

2.1 자동차 번호판 인식

자동차 번호판 인식 기술은^[4] 차량에서 번호판이 부착된 영역을 감지 및 추출하여 번호판을 인식하는 시스템이다. 자동차 번호판 인식 기술은 정확도가 높은 OCR 활용 기술 중 하나이다.

자동차 번호판 인식 기술의 정확도가 특히 높은 이유는 자동차 번호판의 종횡비가 4:1로 일정하게 고정되어 있기 때문이다. 이로 인해 복잡한 장면과 노이즈가 심한 장면에서도 종횡비가 4:1인 사각형을 추출할 수 있기 때문에 비교적 노이즈에 강인한 글자 감지 알고리즘이다. 또한 번호판을 구성하고 있는 글자 요소가 고정되어 있기 때문에 글자 인식의 정확도가 높다. [Fig 1]은 자동차 번호판 인식의 개요이다. 그림에 보듯이 것과 같이 자동차 번호판 인식 시스템에는 먼저, 자동차 번호판이 포함된 이미지가 입력된다. 이후 자동차 번호판이 포함된 이미지를 종횡비 기반의 번호판 추출 기술을 통해 번호판을 추출하고, 추출된 번호판에서 규칙기반의 글자 분류를 수행한다.

예를 들어 자동차 번호판의 글자의 구성요소와 규칙이 존재하여, 그림과 같이 [숫자, 숫자, 글자, 숫자, 숫자, 숫자, 숫자]라고 가정하기 때문에 숫자와 글자를 혼동하지 않는다. 뿐만 아니라 번호판의 포함된 글자의 길이가 고정적이기 때문에 정확한 인식을 돕고 이러한 규칙기반의 인식으로 번호판을 정확



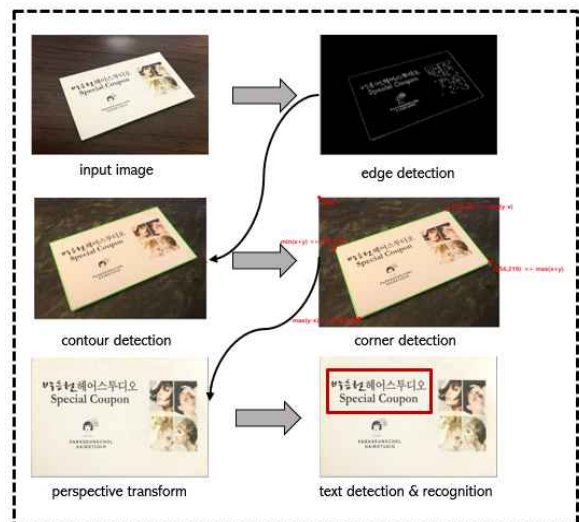
[Fig. 1] License plate recognition pipeline.

하게 인식한다. 이와 같이 번호판 인식 기술은 매우 정형화되어 있다. 하지만 이러한 기술은 자동차 번호판과 같이 특정한 규칙이 있는 환경에서만 응용할 수 있기 때문에 노이즈를 극복한 응용이라고 보기는 어렵다. 우리가 제안하는 알고리즘은 모든 장면에서 노이즈에 강인하기 위하여 규칙기반의 알고리즘이 아닌 CNN을 사용한다.

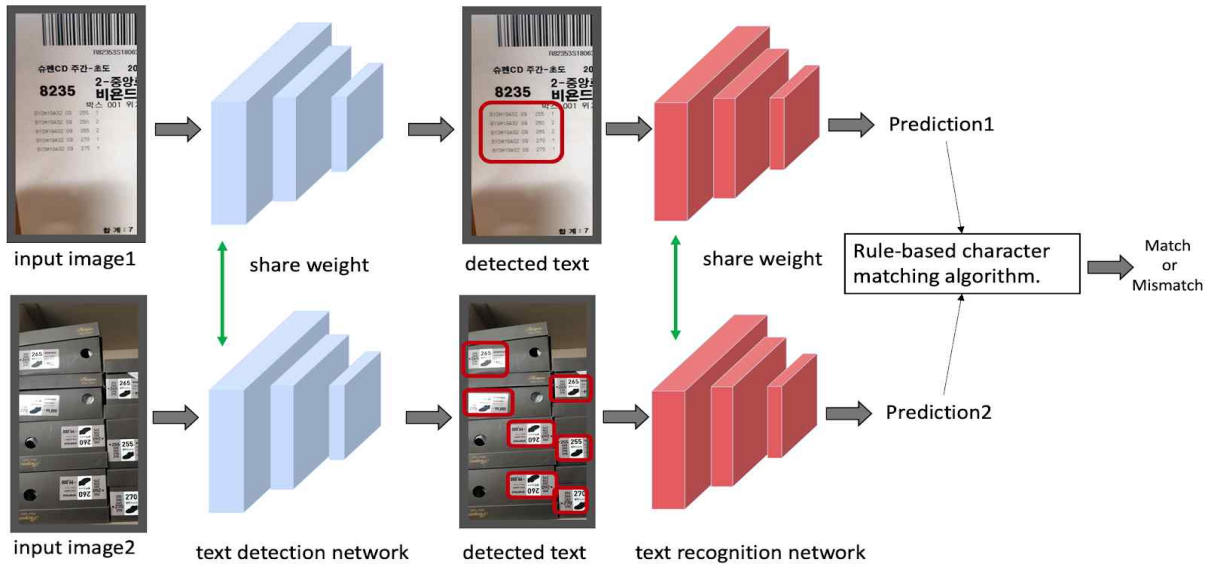
2.2 문서 내 글자 인식

딥러닝 기반의 OCR 기술이 고도로 발달함에 따라 손 글씨, 영수증, 문서 인식 등에서 많은 응용 기술들이 실제 산업에서 활발하게 활용되고 있다.

문서 내 글자 인식 기술은 고정된 배경(노이즈가 없는)에서 글자 검출 및 인식하는 기술이다. 문서 인식 또한 자동차 번호판 인식과 유사하게 배경 노이즈가 존재하지 않기 때문에 상대적으로 쉽게 글자 영역을 추출할 수 있다. 더불어 대부분의 문서는 직사각형이기 때문에 회전된 글자를 원래대로 되돌리는 것은 간단한 투영변환으로 해결할 수 있다. [Fig 2]는 문서



[Fig. 2] Pipeline for character recognition in documents



[Fig 4] OCR-based inventory management application pipe line.

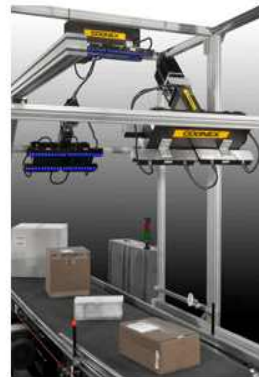
인식 개요이다. 그림에서 보이는 것과 같이 추출된 글자영역을 글자 인식기에 삽입하여 인식한다. 대부분의 문서 내 글자 인식 응용들은 이미지를 그레이 스케일로 변경한 후에 사용하기 때문에 색상 정보에 강인하다. 하지만 문서 내 글자 인식은 노이즈가 거의 존재하지 않는 환경에서만 작동하는 응용기술이기 때문에 현실 노이즈를 극복한 응용기술이라고 보기는 어렵다.

2.3 물류 관리 시스템

물류 관리 시스템은 대규모의 물류창고, 또는 소규모의 판매장 등에서 필수적으로 필요한 시스템이다. 예를 들어 [Fig 3]에 a과 같은 장비들은 대규모의 물류창고에서 활용하는 장비이고, b는 소규모 판매장에서 사용하는 장비이다. 하지만 a,b 두 가지 장비 모두 경제적으로 큰 비용이 소모된다. 따라서 바코드 리더기 기반의 물류관리 시스템을 구성하는 것은 비용적으로 큰 부담이 있다. 또한 a,b 모두 바코드 인식 기반 장비이기 때문에 여러 종류의 바코드중(ITF, CODE39, CODABAR) 장비에서 지원하는 일부 바코드만 인식할 수 있는 문제가 존재한다. 따라서 우리는 OCR기반의 물류관리 시스템을 제안하고 스마트폰과 같은 저렴한 장비에서도 높은 신뢰성과 정확도를 보이는 물류 검수 기계를 제안한다.

3. OCR 기반 물류관리 어플리케이션

우리의 OCR기반 물류관리 어플리케이션은 text detection, recognition으로 구성되어 있다. [Fig 4]는 우리의 알고리즘을



a. Barcode reader for distribution warehouse.

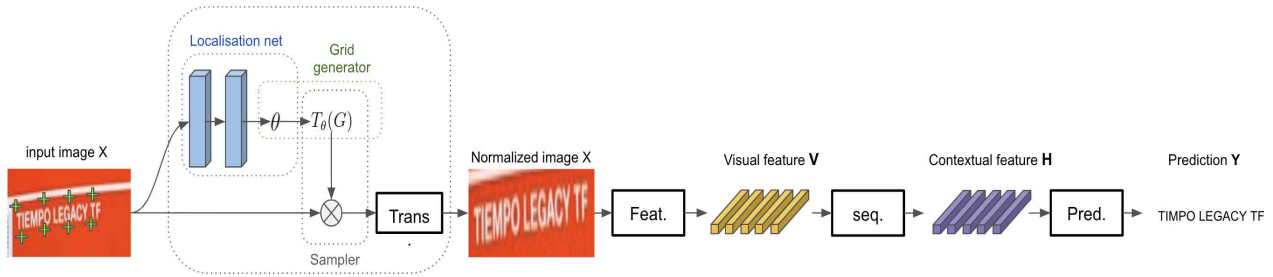


b. Barcode Reader for Sales Stores

[Fig. 3] inventory management device based on barcode recognition.

간략하게 설명하는 그림이다. 그림의 input image1은 한 박스 안에 구성되어 있는 상품들의 목록이고, input image2는 해당 박스 안에 실제로 들어있는 상품이다. 따라서 우리의 목적은 두 장의 사진에서 글자를 개별적으로 추출 및 인식하여, 인식된 두 글자가 서로 일치되도록 매칭 하는 것이다. 만약 두개의 추출된 글자가 정확히 매칭 된다면 물류가 제대로 입고된 것이고 아니라면 제대로 입고되지 않아 정정해야 하는 것이다.

다음으로 3.1 부분에서는 노이즈에 강인한 글자 감지 모델에 대한 설명, 3.2 부분에서는 노이즈에 강인한 글자 인식 모델을 구성하는 부분을 순차적으로 설명하고, 3.3에서는 글자 감지 및 탐지 모델의 조합을 어떻게 설계하였는지에 대해 간략하게 설명한다. 마지막으로 3.4에서는 규칙 기반 글자 매칭 알



[Fig 5] Spatial Transformer Networks pipeline.

고리즘을 설명한다.

3.1 글자 감지

노이즈에 강인한 글자 감지 알고리즘을 구축하는 것은 OCR기반의 물류관리 어플리케이션에서 중요한 요소이다. OCR기반의 물류관리 어플리케이션에서 검수대상들의 글자 영역을 제대로 감지하지 못하면 글자인식기의 성능과는 관련 없이 제대로 된 물류 검수를 할 수 없기 때문이다. 따라서 우리는 현실 노이즈에(각도에 따른 이미지의 공간적 변형)^[5] 강인한 글자 감지기를 만들기 위하여, 적대적 노이즈 회복 학습법을 제안한다.

노이즈 생성기는 글자가 존재하는 영역에만 노이즈를 생성하고, 글자 감지 모델은 정답 레이블과 Cross Entropy를 줄이는 방향으로 학습된다. 하지만 적대적으로 노이즈 생성기는 Cross Entropy를 높이는 쪽으로 학습된다^[6]. 따라서 노이즈 생성기는 학습이 진행되면 될수록 어려운 노이즈를 생성하고, 글자 감지 모델을 어려운 노이즈가 있는 포함된 영상에서 글자를 잘 감지 할 수 있도록 학습된다. [fig 8]은 우리의 적대적 노이즈 회복 학습법의 그림이다. 그림에서 보이는 것과 같이 우리의 적대적 노이즈 회복 학습법은 학습 시에만 사용하는 모듈이기 때문에 실제 응용 할 때 연산량 및 파라미터가 늘어나지 않는다.

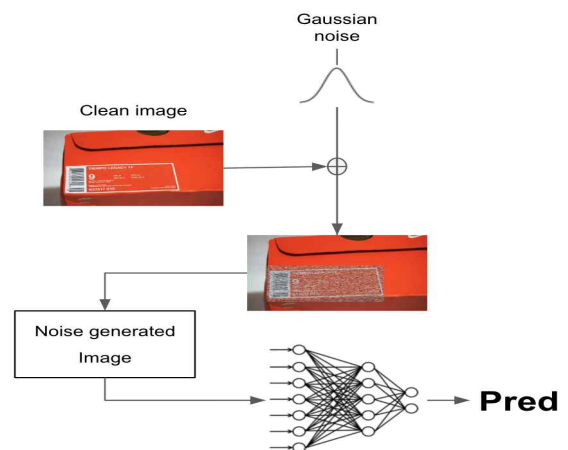
3.2 글자 인식

현대의 글자인식 알고리즘은 매우 고도화되었다^[7]. 하지만 고도화된 알고리즘들은 정면의 글자에서만 유효하다는 한계점이 존재한다. 예를 들어 [fig 5]에 보이는 것과 같이 공간적으로 왜곡된 글자는 글자 인식기에서 제대로 인식하지 못하는 경우가 발생한다. 하지만 이러한 왜곡은 각도, 조명 등과 같이 실제 환경에서 충분히 발생할 수 있는 노이즈이기 때문에 이러한 문제를 해결하지 못하면, 실제 어플리케이션에서 사용되

기에 어려움이 있다. 우리는 이러한 문제점을 해결하기 위하여 STN(spatial transformer networks)^[8]를 사용한다. STN은 [fig 5]에 보이는 것과 같이 왜곡이 있는 이미지를 네트워크가 학습하기 쉬운 방향으로 변환시켜주는 서브 네트워크이다. 우리는 이러한 네트워크를 사용함으로써 각도에 따른 공간적인 변형에 강인한 글자 인식기를 만들 수 있었다.

3.3 글자 감지 및 탐지 모델의 조합

우리는 최적의 글자 감지 및 탐지 모델을 조합하기 위하여, 기반이 되는 네트워크, 알고리즘, 전처리 방법 등 이렇게 크게 3가지 방법으로 조사를 진행 하였다. 먼저, 우리는 기반이 되는 네트워크 조사를 위하여 VGG^[9], ResNet^[10], MobileNetV2^[11], EfficientNet^[12]등의 성능을 조사하였고, 글자 감지 알고리즘은 FOTS^[13], CRAFT^[14], TextBoxes^[15], 글자 인식 알고리즘은 STAR-Net^[16], GRCNN^[17], Rosetta^[18]를 비교하였다. 우리는 앞에서 언급한 모든 방법들을 구조 찾기 범위로 두고 구조들을 조합 하였다. 또한 우리는 구조 찾기 알고리즘을 사용하는 것이 아닌 이 모든 알고리즘을 직접 조합하고 실험



[Fig 9] Adversarial denoise learning pipeline.



[Fig 8] An example of qualitatively showing our text detection in a practical environment.

을 하면서 성능을 비교하였다.

3.4 규칙 기반 글자 매칭 알고리즘

CNN을 활용하여 글자를 감지하고 인식하여 추출된 글자를 물류검수에 활용하기 위해서는 먼저 글자를 매칭 해야 한다. 예를 들어 [fig 4]에 image1처럼 검수 목록과 image2에 검수 대상의 순서는 일정하지 않다. 그렇기 때문에 우리는 검수 목록을 리스트에 저장하고, 검수 대상에서 추출된 글자가 검수 목록에 있으면 해당 목록을 검수 목록 리스트에서 하나씩 삭제하는 방식을 사용하였다. 최종적으로 알고리즘은 검수 목록 리스트가 0이 되면 물류가 정확하게 입고되었다고 알리고, 검수 목록 리스트가 0이 아니면 물류가 잘못 입고되었다고 사용자에게 알리는 방식으로 동작한다.

4. 실험 결과

우리는 글자 인식 및 탐지 모델을 학습시키기 위하여 [Fig 7]에 보이는 것과 같이, MJSynth (MJ)^[19] 와 SynthText (ST)^[20]를 사용하였다. MJ 데이터셋은 10Gb(200만장) 규모의 대규모 학습 데이터셋이며, [Fig 7]에 왼쪽과 같이 왜곡이 없고 현실에서 실제로 존재하는 이미지이다. 이는 ST노이즈에 강인한 모델을 위하여 글자에 노이즈를 섞은 41Gb(800만장) 규모의 가상의 이미지이다. 마지막으로 우리가 학습시킨 모델을 평가

하기 위하여 모든 실험에서 우리가 직접 수집한 물류사진 500장을 통해 성능을 검증하였다.

4.1 글자 감지 및 탐지 모델의 조합

우선적으로 우리의 OCR기반 물류관리 어플리케이션에 최적화된 네트워크를 구조를 선정하기 위하여 3.3절에서 언급한 구조 찾기 범위에서 실험을 하였다. [Table 1]은 최적의 구조 찾기 실험 결과에 대한 표이다. 표에 보이는 것과 같이 기존 성능이 높다고 알려진 알고리즘들이 실제 환경에서는 성능이 높지 않을 수 있다는 것을 실험적으로 보였다. 따라서 [Table 1] 실험 결과에 기반하여, ResNet50, CRAFT, Rosetta 조합을 채택하였다. [Table 1]의 실험 결과는 기존 알고리즘들의 공정한 성능 비교를 위하여 적대적 노이즈 회복 알고리즘을 적용하지 않았으며, 4.1 이 후에 나오는 모든 실험들은 ResNet50, CRAFT, Rosetta(RCR) 조합으로 실험을 수행한다.



a. MJSynth(MJ)



b. SynthText(ST)

[Fig 7] MJ, ST dataset samples.

[Table 1] An experiment to find the optimal architecture in the various architecture search ranges designated by us. (All evaluations were evaluated on the IC15 dataset.)

Backbone	text detection	text recognition	Acc	Backbone	text detection	text recognition	Acc
VGG	FOTS	STAR-Net	76.1	MobileNetV2	FOTS	STAR-Net	72.2
VGG	FOTS	GRCNN	74.5	MobileNetV2	FOTS	GRCNN	71.1
VGG	FOTS	Rosetta	76.6	MobileNetV2	FOTS	Rosetta	75.2
VGG	CRAFT	STAR-Net	78.1	MobileNetV2	CRAFT	STAR-Net	76.1
VGG	CRAFT	GRCNN	77.2	MobileNetV2	CRAFT	GRCNN	75.9
VGG	CRAFT	Rosetta	76.8	MobileNetV2	CRAFT	Rosetta	77.2
VGG	TextBoxes	STAR-Net	74.7	MobileNetV2	TextBoxes	STAR-Net	74.1
VGG	TextBoxes	GRCNN	75.2	MobileNetV2	TextBoxes	GRCNN	75.2
VGG	TextBoxes	Rosetta	76.5	MobileNetV2	TextBoxes	Rosetta	77.1
ResNet	FOTS	STAR-Net	71.1	EfficientNetB0	FOTS	STAR-Net	69.5
ResNet	FOTS	GRCNN	74.7	EfficientNetB0	FOTS	GRCNN	69.9
ResNet	FOTS	Rosetta	75.4	EfficientNetB0	FOTS	Rosetta	72.1
ResNet	CRAFT	STAR-Net	77.8	EfficientNetB0	CRAFT	STAR-Net	75.4
ResNet	CRAFT	GRCNN	76.2	EfficientNetB0	CRAFT	GRCNN	76.1
ResNet	CRAFT	Rosetta	78.9	EfficientNetB0	CRAFT	Rosetta	77.5
ResNet	TextBoxes	STAR-Net	75.2	EfficientNetB0	TextBoxes	STAR-Net	75.9
ResNet	TextBoxes	GRCNN	75.8	EfficientNetB0	TextBoxes	GRCNN	76.1
ResNet	TextBoxes	Rosetta	76.9	EfficientNetB0	TextBoxes	Rosetta	77.9

[Table 2] Experimental results of adversarial denoise learning from the practical environment data set we collected.

method	Noise	ACC
RCR	None	93.9
RCR	full region noise	92.4
RCR	text region noise	89.1
RCR+Adv	full region noise	95.8
RCR+Adv	text region noise	97.2

4.2 글자 감지 및 인식

노이즈에 강인한 글자 감지 알고리즘을 완성하기 위하여 우리가 제안한 적대적 노이즈 회복 학습 기법을 실험했다. 우리가 제안한 적대적 노이즈 회복 학습법을 기존의 데이터 증식 효과와 공정하게 비교하기 위하여 [Table 2]에 보이는 것과 같이 노이즈만 추가한 것, 노이즈와 우리의 학습법을 추가한 것 마지막으로 글자 영역에만 노이즈를 추가한 것, 전체 영역의 노이즈를 추가한 것을 비교했다. 또한 우리의 어플리케이션에서 잘 작동 해야 함으로, 우리가 수집한 실무 데이터셋 500장에서 평가 하였다. [Table 2]에 보이는 것과 같이 글자영역에만 노이즈를 추가하고, 그 노이즈를 회복하는 학습 기법을 사용하면 기존 방법에 비하여 약 4%의 성능 향상이 있음을 보였다.

[Table 3] OCR-based inventory management application Final performance.

method	Noise	ACC
RCR	None	97.9
RCR+Adv	text region noise	99.7

4.3 물류관리 정확도

최종적으로 우리의 알고리즘의 목표인 OCR 기반의 물류관리 어플리케이션에서 정확도를 확인하였다. [Table 3]는 우리의 알고리즘을 우리가 수집한 데이터셋에서 검증한 결과이다. 실무 환경에서는 아주 정확한 글자 비교가 필요 없기 때문에 (특정 키워드가 동일 한 것이 중요하기 때문에 노이즈처럼 존재하는 글자는 비교할 필요가 없음.) OCR 기반의 물류관리 어플리케이션의 검수 결과만으로 정확도를 측정하였다. [Table 3]에 보이는 것과 같이 우리의 알고리즘은 99.7성능을 달성하였다.

5. 결 론

우리는 매우 직관적이고 간단한 방법으로 노이즈에 강인한 학습법 제안하여 현실에서 사용할 수 있는 OCR 알고리즘 개발하였다. 또한 이러한 모델을 기반으로 OCR기반의 물류관리 어플리케이션을 제안하여 규모가 작은 유통매장에서 사용

가능한 신뢰성 있는 물류 관리 시스템을 구축하였고, 이를 통해 상대적으로 규모가 작은 유통업자의 노동 강도를 줄일 수 있는 방법을 찾아낼 수 있었다. 마지막으로 우리는 OCR기반의 물류관리 어플리케이션이 전문 지식이 부족하거나 신체적으로 불편한 사람도 충분히 물류 관리를 할 수 있게 만듦으로써 장애인 일자리 창출의 기회가 됨을 기대한다.

References

- [1] LIU, Wei, et al. STAR-Net: A SpaTial Attention Residue Network for Scene Text Recognition. In: BMVC. 2016. p. 7.
- [2] BAEK, Jeonghun, et al. What is wrong with scene text recognition model comparisons? dataset and model analysis. In: Proceedings of the IEEE International Conference on Computer Vision. 2019. p. 4715-4723.
- [3] GOODFELLOW, Ian, et al. Generative adversarial nets. In: Advances in neural information processing systems. 2014. p. 2672-2680.
- [4] ZHERZDEV, Sergey; GRUZDEV, Alexey. Lpnet: License plate recognition via deep neural networks. arXiv preprint arXiv:1806.10447, 2018.
- [5] HENDRYCKS, Dan; DIETTERICH, Thomas. Benchmarking neural network robustness to common corruptions and perturbations. arXiv preprint arXiv:1903.12261, 2019.
- [6] RUSAK, Evgenia, et al. Increasing the robustness of DNNs against image corruptions by playing the Game of Noise. arXiv preprint arXiv:2001.06057, 2020.
- [7] WOJNA, Zbigniew, et al. Attention-based extraction of structured information from street view imagery. In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2017. p. 844-850.
- [8] JADERBERG, Max, et al. Spatial transformer networks. In: Advances in neural information processing systems. 2015. p. 2017-2025.
- [9] SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
- [10] HE, Kaiming, et al. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 770-778.
- [11] SANDLER, Mark, et al. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. p. 4510-4520.
- [12] TAN, Mingxing; LE, Quoc V. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946, 2019.
- [13] LIU, Xuebo, et al. Fots: Fast oriented text spotting with a unified network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. p. 5676-5685.
- [14] BAEK, Youngmin, et al. Character region awareness for text detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019. p. 9365-9374.
- [15] LIAO, Minghui, et al. Textboxes: A fast text detector with a single deep neural network. arXiv preprint arXiv:1611.06779, 2016.
- [16] NGIAM, Jiquan, et al. Starnet: Targeted computation for object detection in point clouds. arXiv preprint arXiv:1908.11069, 2019.
- [17] Jianfeng Wang and Xiaolin Hu. Gated recurrent convolution neural network for ocr. In NIPS, pages 334-343, 2017
- [18] BORISYUK, Fedor; GORDO, Albert; SIVAKUMAR, Viswanath. Rosetta: Large scale system for text detection and recognition in images. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018. p. 71-79.
- [19] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman. Synthetic data and artificial neural networks for natural scenetext recognition. In Workshop on Deep Learning, NIPS, 2014.
- [20] A. Gupta, A. Vedaldi, and A. Zisserman. Synthetic data for text localisation in natural images. In CVPR, 2016.

저자약력



서민석

2019 한밭대학교 정보통신공학과 (학사)

2020~현재 한밭대학교 정보통신대학원 (석사)

관심분야: Robotics, Vision Programing, Deep Learning



김대한

2015~현재 한밭대학교 정보통신공학과 (학사)

관심분야: Super Resolution, Domain Adaptation, Computer vision, Deep Learning



강 혜 윤

2017~현재 한밭대학교 정보통신공학과
(학사)

관심분야: Super Resolution, Computer vision, Deep
Learning



최 동 길

2005 한양대학교 전자컴퓨터공학부(학
사)

2007 한양대학교 전자전기제어계측공
학과(석사)

2016 KAIST 로봇공학학제전공(박사)

2018 KAIST 정보전자연구소 박사 후 연
구원

2018~현재 한밭대학교 정보통신공학과
조교수

관심분야: Robot Vision, Sensor Fusion, Autonomous Robot
System, Artificial Intelligence